Faculty and Researcher Publications

1963-08

Statistical Prediction Methods for North American Anticyclones

Martin, F.L.

Journal of Applied Meteorology, Volume 2, pp. 508-516, August 1963.
http://hdl.handle.net/10945/45783

Statistical Prediction Methods for North American Anticyclones[1]

F. L. MARTIN, J. R. BORSTING, F. J. STECKBECK AND A. H. MANHARD, JR.

U. S. Naval Postgraduate School, Monterey, Calif.

(Manuscript received 17 September 1962, in revised form 18 March 1963)

ABSTRACT

Statistical prediction methods of forecasting the 24-hr movement and change of central pressure of North American winter anticyclones are developed utilizing multiple linear regression analysis. The three predictands obtained are the 24-hr change in central pressure and the eastward and northward displacements. This report is similar in some respects to one conducted by Veigas and Ostby relating to U. S. east coast cyclones.

The dependent data consist of observations of 150 anticyclones. A moving coordinate system is employed: the predictor information is measured at certain predetermined grid points relative to the system center rather than at fixed geographical positions.

Readily measured meteorological parameters are selected for the input data. These include point values of surface pressure, surface temperature, 500-mb height and their 24-hr changes; and the arithmetic mean temperature of the layer from the surface to the 500-mb level.

Several different sets of regression equations are obtained for each predictand by slight modifications in the predictor-selection criteria. These regression equations are tested on a sample of 50 independent equations, and the per cent reductions of variance resulting from each method are compared.

1. Introduction

A recent paper by Veigas and Ostby (1963) dealt statistically with the problem of the 24-hr movement and deepening of U. S. east coast cyclones in winter. Veigas and Ostby considered two separate approaches as far as input data were concerned: one ("the complex") approach employed dynamic predictors, whereas the second dealt simply with *point values* of meteorological parameters such as pressure, temperature, contour height, etc., from the sea-level and 500-mb charts. There was little, if any, additional predictive power in the first method over the second. However the data are much easier to read from maps in the latter method.

In the present paper, the second Ostby-Veigas approach is generalized to the consideration of North American winter anticyclones, which lie east of the Rockies (see Fig. 1). The mathematical procedure involves a stepwise multiple regression analysis using the BIMD 09 program as adapted to the CDC 1604 computer.

2. Data sources

The anticyclonic cases used in this study were taken from the Daily Series Synoptic Weather Maps analyzed by the National Weather Analysis Center for both the surface and 500-mb levels. These analyses are drawn using polar stereographic projections true at latitude 60N and they have been utilized for compiling both the dependent and independent test data. The sea-level maps were analyzed for observations taken at 1230 GCT, while the 500-mb observations were taken at 1500 GCT.

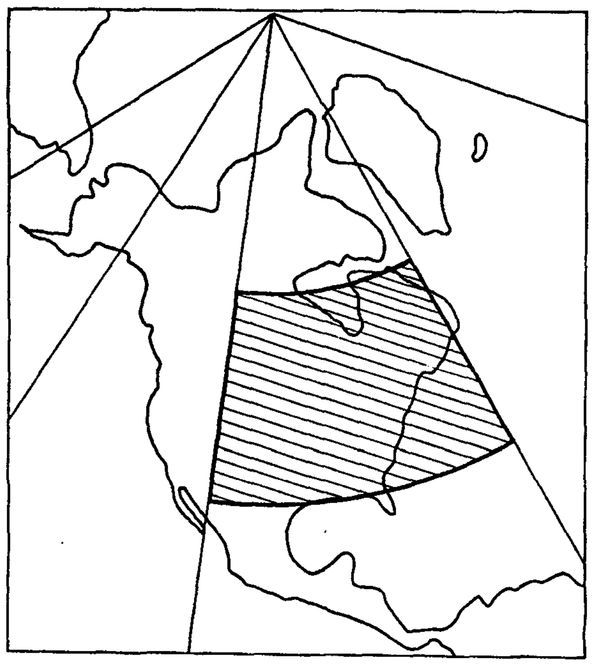

FIG. 1. The forecast area.

[1] This research was supported in part by the Office of Naval Research.

Anticyclones for the dependent sample were selected from a geographic area extending from longitude 60 to 110W, between latitudes 30 and 60 N. This area is depicted in Fig. 1. The period covered extended from October through March for the years 1953 through 1957 (1955 excluded). The final regression equations developed in this study should be applied only from October through March for those anticyclones which lie in the area defined above, henceforth called the forecast area.

In order that an anticyclone be included in the sample, it had to be characterized by at least one closed isobar at a 5-mb interval. A further requirement was that the closed system had to remain within the forecast area during the succeeding 24 hours. If the system met both of the above requirements for an additional 24-hr period, it was included again in the sample as a second case. In order to eliminate undue bias, no single system was included a third time. These precautions are similar to those taken by Ostby and Veigas in their study of cyclones.

Based upon the above criteria, 150 cases were selected for the dependent sample data. The frequency distribution for this data by month and year is as shown in Table 1. The number of cases examined varied from as many as eleven in February of 1953 to as few as one in February of 1957. The number of cases was a maximum during March (an average of 8.2) and a minimum in February (an average of 6.0).

The year 1955 was not included in the dependent sample in order that the data for this period might be retained for use in the independent test.

Table 1. Dependent sample frequency distribution by months and years.

	Oct	Nov	Dec	Jan	Feb	Mar	Total
1953	—	—	—	9	11	9	29
1953–54	—	9	7	8	5	10	39
1954–55	5	10	10	—	—	—	25
1955–56	—	—	—	6	7	9	22
1956–57	8	6	8	7	1	5	35
Total	13	25	25	30	24	33	150

3. Selection of possible predictors

The use of the electronic computer makes it possible to analyze statistically a large number of possible predictors with great speed. A screening program known as BIMD 09 has been adapted to the CDC 1604 electronic computer for the purpose of carrying out the analysis of the reduction of variance necessary for the selection of predictors. This program, which is available in the Computer Center of the U. S. Naval Postgraduate School, is a modification of one originally written by M. A. Efroymson (1960).

When one utilizes such a technique, he must specify precisely the initial set of possible predictors as well as the selection criteria. The consideration of *a priori* known climatological, synoptic or dynamic influences may be a factor in the specification of possible predictors, but this was not attempted here. A square grid map overlay (Fig. 2) containing 100 points equally spaced in both the x and y directions was devised for the purpose of extracting specific grid-point information. The spacing between adjacent grid points was set equal to that of a 5-degree latitude interval at latitude 45N. This constant spacing is, of course, equal to 300 *true nautical miles* on the earth's surface, but is somewhat more (less) than 300 "map" nautical miles in higher (lower) latitudes, when one considers the map-scale factor

$$m = \frac{1.86603}{1+\sin L}$$

where L is the latitude. However, all units used for displacement of anticyclones are in terms of true-earth distance, so that no consideration need be given to the map factor, either in regard to displacement or any of the predictors. It should be noted that this grid differs from that of Veigas and Ostby in being square rather than based on a latitude-longitude net. However, the grid is similar to theirs in that it is a *moving-coordinate* grid, in which grid point 54 is always located over the sea-level high, with the y-axis aligned along the meridian through point 54. Smaller sets of grid points have been used for recording certain other parameters, but these represent merely sub-sets of the basic grid and utilize the same numbering system.

Features of the present and past 24-hour pressure and temperature fields are embodied to furnish the majority of possible predictors. The sea-level and the 500-mb

Table 2. Possible predictors examined in the dependent data, and the units employed.

Possible predictors	Units	Symbol	Number of possible predictors
Sea-level pressure	mb	p	54
24-hr surface pressure change	mb	Δp	13
Surface temperature	°C	T	1
24-hr surface temperature change	°C	ΔT	1
Location of surface anticyclone	Deg. lat., long.	L, λ	2
Anticyclone intensity	mb	I	1
500-mb height*	tens of feet	z	27
24-hr 500-mb height change	tens of feet	Δz	27
Mean temperature, surface to 500 mb	°C	$\bar{T}$	19
			145

* The ten-thousand digit is not included in the z-predictors; for example, 18770 is recorded as 877.

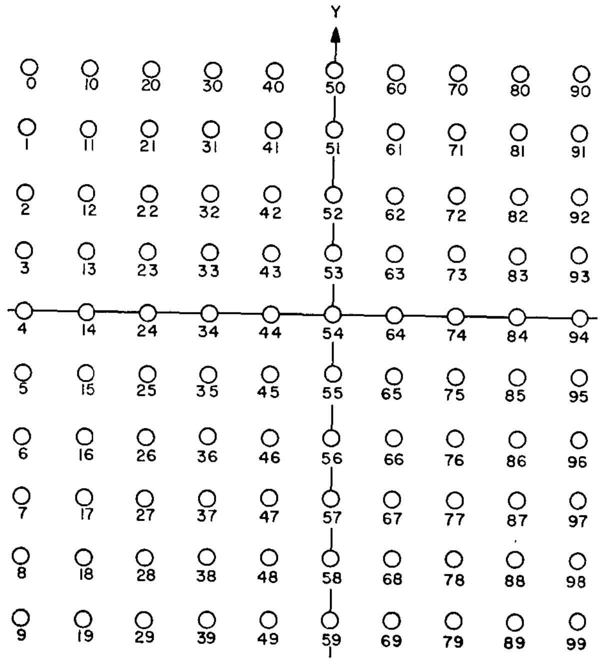

FIG. 2. The basic grid. The y-axis is aligned along a meridian through the sea-level anticyclone. The x-axis joins the grid points 4, 14, $\cdots$94.

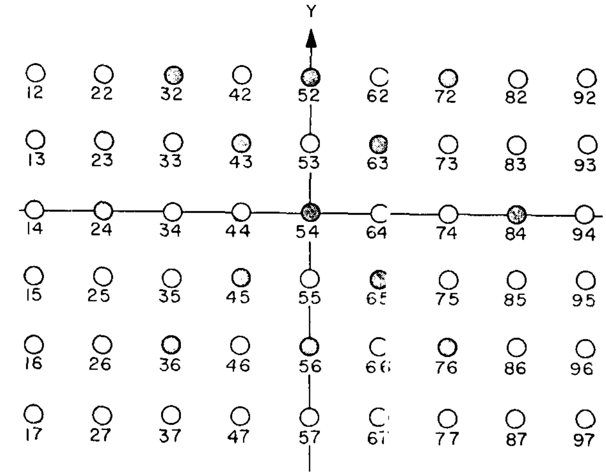

FIG. 3. Grid points used to obtain the fields of sea-level pressure and of 24-hour pressure tendency. Open circles designate grid points for the pressure field, while the shaded circles are used for the pressure-tendency field.

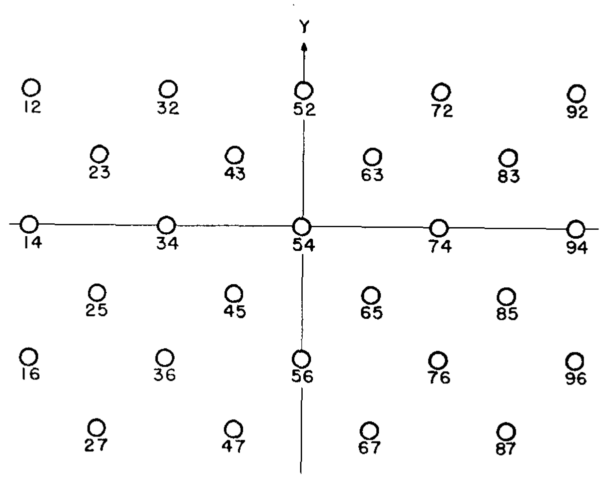

FIG. 4. Grid points used to obtain 500-mb heights.

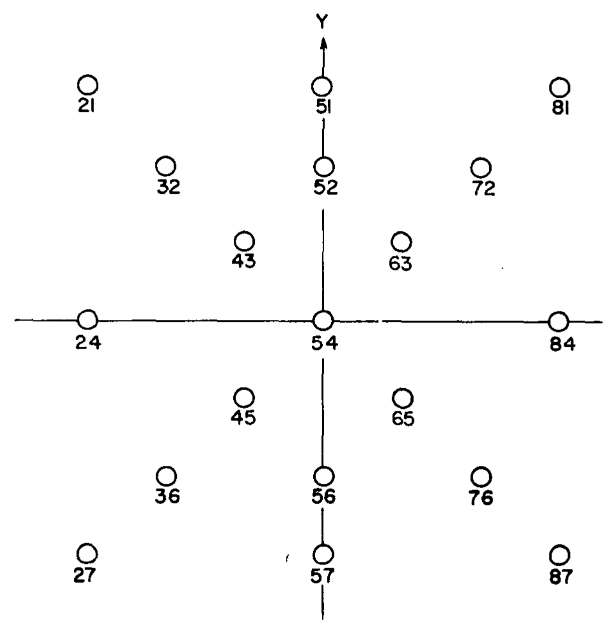

FIG. 5. Grid points used to obtain sea-level and 500-mb temperatures.

contour charts were employed to obtain all predictors. A description of the types of independent variables tested is given below under the headings (a) sea-level variables and (b) upper-air variables. Table 2 presents a summary of these types.

a) Sea-level variables. Point values of the sea-level pressure at time t_0 were recorded and included as possible predictors. The grid overlay was positioned on the sea-level chart with grid point 54 located over the system center. Values of sea-level pressure, read to the nearest whole millibar, were then recorded at the total of 54 grid points shown in Fig. 3. These values then specified the sea-level pressure field around the system center at observation time (t_0).

With the grid in the same relative position, point values of surface pressure 24 hours prior to the observation time were recorded at the 13 shaded grid points, also contained in Fig. 3. These values were then used to specify 13 values of the 24-hr pressure-change field, which were also included in the list of possible predictors.

The *initial latitude, longitude* and *intensity* of the sea-level anticyclone at t_0 were also recorded for use as possible *predictors*. The intensity of the anticyclone was defined by subtracting the space-averaged sea-level pressure at grid-point 54 (obtained using the arithmetic mean of the values at the four surrounding points 43, 45, 63, 65) from the actual sea-level value at this point.

b) Upper air variables. The values of the 500-mb parameters at t_0 were obtained from the 500-mb chart by positioning grid point 54 of the appropriate grid over the surface anticyclone center at t_0. Values of 500-mb height were then recorded at 27 grid points as indicated in Fig. 4. These 27 values were also included in the list of possible predictors.

Using the thermal pattern implied by a grid plot of the mean 1000–500-mb temperature, one may use the dynamical concepts of Sutcliffe (1947) to deduce at least qualitative approximations to the upper-level convergence, and the resulting evolution of the synoptic flow pattern of middle and upper layers of the troposphere. According to Sutcliffe, baroclinic development and movement of a sea-level anticyclone are dynamically related to its position relative to the 1000–500 mb mean-temperature field. In order to obtain a representative mean temperature field for this layer, the surface and 500-mb temperatures were recorded from synoptic charts at the 19 grid points shown in Fig. 5. From these values, the *arithmetic mean temperatures* of the layer were calculated and also included in the list of possible predictors.

It should be noted that we are seeking linear regressions, whereas Sutcliffe's method of computing advection of vorticity or of thermal vorticity involves non-linear wind and vorticity fields. Any linear regressions obtained here which are consistent with his view cannot be expected to give the degree of significance that would exist using appropriate non-linear regressions. However significant linear regressions could be consistent with Sutcliffe's philosophy applied to a *linearized model* based upon the data.

The 145 possible predictors and three predictants are summarized in Table 2 both as to type and the units employed. For each of these variables, a dependent sample of 150 separate cases was recorded and punched onto IBM cards. The data on the cards was then transferred to magnetic tape for convenient storage in readily accessible form for the BIMD 09 program.

4. The stepwise multiple regression analysis

The reduction of the 145 possible predictors was accomplished by employing a *stepwise* regression procedure. Such stepwise procedures have been extensively discussed in recent years by Miller[2] (1962) and Efroymson (1960). The stepwise procedure used in this study is of the forward type. The regression equations were developed by adding one independent variable at a time at each step of the analysis of variance reduction. In each case, the variable to be added is that which gives the highest partial correlation coefficient with the dependent variable after the effects of all previously selected variables have been removed. The selection procedure continues until it is stopped by some predetermined "cutoff" criterion.

The exact formulation of cutoff rules in stepwise regression procedures has recently received considerable attention. For example, Efroymson (1960) proposed the use of critical F values both to *accept* and to *remove* variables in the regression analysis. Efroymson's critical F values are treated as constants throughout all the partial regression steps. On the other hand, Miller (1962, p. 45) proposed a variable cutoff for *each* independent variable X_S added in a stepwise manner to the regression equation

$$y' = a_0 + a_1X_1 + a_2X_2 + \cdots a_iX_i + \cdots + a_SX_S. \quad (1)$$

His test employs "a critical F value whose probability level is a function of the number of possible predictors P" and "of the Sth selected predictor." The proposed critical F value corresponds to the tabulated value of F at a probability level $\alpha = \alpha^*/(P-S+1)$ with 1 and $N-S-1$ degrees of freedom, where N is the size of the dependent sample. In the present study, we have used: $N=150$, a (maximum) number of possible predictors $P=145$, and $\alpha^*=0.05$. We therefore have $\alpha=0.05/145$ as the initial confidence level. This gave an initial critical F value of 12.5. In this study, the total number S of predictors selected by Miller's criteria was small, and the critical F value remained essentially constant during the test process. Hence we employed a constant $F=12.5$ in the successive steps of Miller's selection procedure.

In order to insure that the BIMD 09 regression analysis may duplicate Miller's system, we employed a lower critical F-value of zero. In this connection, any lower critical F value designed (in principle) to remove variables from the BIMD 09 program will be called a "cutout" value. This term should be contrasted to the upper critical value, which has been called a *cutoff*.

As has already been pointed out, Efroymson's method maintains constant values of F for cutoff as well for cutout, in formulating *decision rules* for predictor selection. When Efroymson's method is used with an F-cutoff of 12.5 and a cutout of zero, we considered it to give identical results to Miller's method. However, in order to consider the effect of varying Miller's selection criteria, we used Efroymson's program with a relatively non-stringent F-cutoff of 4.0, and in addition an intermediate cutoff of 8.0. We also considered the effect of varying the F-cutout level: F-cutouts of zero

[2] Actually Miller's statistical contributions in this field have been very extensive since 1958, but his 1962 publication puts forth an elegant summary of his previous work.

TABLE 3. Summary of the stepwise regression analysis for the south-north displacement N', using the dependent sample. (The units are as listed in Table 2).

Predictor selected	F level upon entry	S_y after entry	Cumulative $P.R.$	$P.R.$ upon entry	Corresponding coefficient in Eq (1) I	II	III
z_{74}	24.7215	228.5312	13.75	13.75	2.0648	2.4122	1.8402
z_{34}	31.6323	208.0158	28.53	14.78	−3.4021	−2.0939	−2.2783
$\bar{T}_{56}$	18.0628	196.9019	35.96	7.43	13.8002	22.1224	23.4010
$\bar{T}_{43}$	8.3023	192.1551	39.01	3.05		−15.8355	−14.8125
p_{22}	16.9421	182.3900	45.05	6.04		−6.4254	−6.8596
p_{97}	11.0965	176.3137	48.65	3.60		−8.1446	−9.8432
p_{85}	4.3843	174.2636	49.84	1.19			5.5593
Const. term					1225.8887	14646.7767	11578.4606
s_y 246.0512 n mi.							
F value, by Eq (2)					27.51	22.58	20.16

TABLE 4. Summary of the stepwise regression analysis for west-east displacement E', using the dependent sample. Section (a) corresponds to a zero F-cutout and (b) to an F-cutout of 4.0.

	Predictor selected	F level upon entry	S_y after entry	Cumulative $P.R.$	$P.R.$ upon entry	Corresponding coefficient in Eq (1) I	II	III
	z_{67}	21.6097	237.1328	12.15	12.15	0.9322	1.4192	0.68893
	z_{52}	31.1193	216.1559	27.01	14.86	−2.2379	−1.9750	−2.0143
	z_{65}	13.1229	207.7588	32.57	5.56	2.1056	1.90295	2.59620
	p_{34}	8.5300	202.5998	35.87	3.30		−7.7851	−5.0627
	Δp_{63}	8.9738	197.2491	39.22	3.35		7.5984	7.7368
(a)	λ	10.3095	191.1665	42.91	3.69		6.1081	8.4005
	Δz_{25}	7.4105	187.0205	45.36	2.45			−1.7158
	p_{84}	5.7453	183.9718	46.55	1.19			−15.6533
	p_{85}	7.1016	180.1159	49.32	2.77			18.9194
	p_{86}	4.1711	178.1101	50.44	1.12			−8.6364
	Const. term (nm)					−422.0655	6381.2373	8928.9790
	z_{67}	21.6097	237.1328	12.15	12.15	0.9322		
	z_{52}	31.1193	216.1559	27.01	14.86	−2.2379	−2.2281	−2.2601
	z_{65}	13.1229	207.7588	32.57	5.56	2.1056	3.1003	3.0343
	Δz_{25}	8.7252	204.6491	34.57	2.00		−2.3519	−2.6000
(b)	λ	10.7947	198.0252	38.74	4.17		8.8920	10.4411
	Δp_{63}	11.8884	190.8190	43.07	4.33		7.6485	6.3267
	p_{84}	8.5280	186.0237	45.94	2.87		6.5089	−16.2598
	p_{85}	5.9933	182.8148	47.79	1.85			10.4447
	p_{33}	5.4654	179.9673	49.40	1.61			−7.8012
	p_{74}	4.1099	177.9872	50.51	1.11			7.6060
	Const. term (nm)					−482.0655	5242.9281	4655.2111
	s_y 253.0019 n mi							
	F level, section (a)					32.51	17.61	14.15
	F level, section (b)					23.51	20.25	15.87

TABLE 5. Summary of the stepwise regression analysis for change of central pressure P', using the dependent sample.

Predictor	F level upon entry	S_y after entry	Cumulative $P.R.$	$P.R.$ upon entry	Corresponding coefficient in Eq (1) I	II	III
p_{65}	23.5363	3.9805	13.14	13.14	−0.2306	−0.2110	−0.2114
Δz_{45}	15.9340	3.7937	21.10	7.96	−0.1609	0.04065	0.0400
z_{52}	12.9043	3.6488	27.01	5.91	0.0603	−0.0144	−0.0203
p_{25}	10.2031	3.5390	31.34	4.33		−0.1295	−0.1200
Δp_{72}	10.3812	3.4298	35.51	4.71		0.0871	0.1012
Δp_{43}	11.9076	3.3068	40.05	4.54		0.1401	0.1343
Δz_{92}	5.7238	3.2535	41.97	1.92			0.0234
z_{74}	5.0348	3.2082	43.57	1.60			0.0134
Δp_{76}	5.1083	3.1625	45.17	1.60			0.0776
Const. term					247.5197	358.5254	343.1242
s_y 4.2709 mb							
F value, by Eq (2)					18.01	15.92	12.82

and four were used for removal of variables from the prediction equations.

We therefore generated all the prediction equations of form (1) using the three constant cutoff values

I	II	III
12.5	8.0	4.0

and, in association with each of these, employed constant F-cutouts of zero and 4.0. This actually leads to six sets of selection criteria, although in only one of our predictands was any major difference introduced by the variation in the cutout (Table 4).

In the prediction equation (1), y' denotes in turn:

N' = predicted 24-hour south-north displacement in nautical miles of the sea-level anticyclone; positive values indicate a northward displacement,
E' = predicted 24-hour west-east displacement (nautical miles) of the anticyclone; positive values indicate eastward motion;
P' = predicted 24-hour change (in millibars) in central pressure. A positive value indicates a pressure rise.

Tables 3, 4 and 5 list the predictors selected for N', E' and P', respectively, using the dependent-data sample only. In these tables, the coefficients a_i corresponding to the particular predictors X_i are also listed. Note that the particular value of the constant cutoff has a definite influence upon the number of predictors selected and upon the resulting best-fit coefficients. The coefficients determined for the various predictors under the selection criteria I, II and III are clearly displayed in Tables 3, 4 and 5, and they are uniquely associated with the predictor listed, just to the left, in the first column of these tables.

The substitution of the *cutout criterion* 4.0 for zero in combination with cutoffs I, II and III introduces no change in the predictors found for N' and P' (see Tables 3 and 5). However with the *higher cutout*, and with cutoff criteria II and III (but not I), the variable z_{67} selected at the first step of the selection was *removed* after the third step for E' (see Table 4b). This is due to fact that z_{67}, when considered now in combination with z_{52} and z_{65}, failed to pass the test of an F-cutout of four. However a constant cutout of zero could not have rejected any previously selected variable. As a result of the removal of z_{67}, the regression equations implicit in Table 4b are different from those of Table 4a.

The form of the F function used to compute the significance of the prediction equations at the kth step in a multiple regression analysis is given in Eq (2) below, after Anderson (1960, p. 89):

$$F_k(k, n-k-1) = \frac{R_k^2}{1-R_k^2}\,\frac{N-k-1}{k}. \tag{2}$$

Here R_k is the multiple linear correlation after k predictors have been selected from the dependent sample, and N is the dependent sample size. Eq (2) was then employed to determine the F levels of the final prediction equations for cases I, II, III of Tables 3, 4, 5. These final F levels are in each case considerably in excess of the critical F at a 99 per cent significance level.

The standard error of estimate S_y, which appears at each step in Tables 3, 4 and 5, may be expressed in terms of the multiple correlation coefficient R_k through

$$S_y = s_y(1-R_k^2)^{\frac{1}{2}}, \tag{3}$$

where s_y is the standard deviation of the dependent variable. When (3) is solved for R_k^2, we obtain

$$R_k^2 = 1 - \frac{S_y^2}{s_y^2} = \text{per cent reduction of variance.} \tag{4}$$

Eq (4) gives the *cumulative* per cent reduction of variance, which is usually abbreviated to "*P.R.* variance" or simply "*P.R.*" Subtraction of two successive cumulative *P.R.* values then gives "*P.R.* upon entry" (see also Tables 3, 4, 5), i.e., the *P.R.* attributable to the predictor just entered.

The value of "F level upon entry" at the kth step may be obtained from the relation

$$F_k(1, N-k-1) = \frac{R_k^2 - R_{k-1}^2}{1-R_k^2}(N\text{-}k\text{-}1), \tag{5}$$

where R_k is the multiple correlation coefficient after the kth independent variable has been added to Eq (1). Eq (5) is identical to that employed by Klein *et al.* (1959). The BIMD 09 program computes the F_k levels in accordance with (5), and they are included in Tables 3, 4 and 5 for each predictor selected.

The standard error of estimate and the *P.R.* variance of the three predictands in the *independent sample* were computed using the prediction equations summarized in Tables 3, 4 and 5 applied to the independent data. For convenient comparison, the corresponding *P.R.* variances for the dependent sample has been repeated in Table 6. The subdivisions I, II and III refer to the cutoff value employed. Numbers in parentheses in the E' column indicates that an F-cutout of four was used, while *all other P.R.* variances refer to an F-cutout of zero.[3] A detailed discussion of the various prediction methods is presented in Section 7, where special attention is focussed upon Table 6.

6. Data limitations

The term "data" employed here refers to observations of the type listed in Table 2 recorded from the

[3] The F-cutout may also be considered to be four since no change was caused by the substitution of zero by four.

TABLE 6. Summary of the application of the prediction equations to the independent data. (The units for S_y and s are n mi for N' and E', and mb for P'.)

	N'			E'			P'		
	I	II	III	I	II	III	I	II	III
S_y	222.60	217.01	221.12	185.96	195.97 (185.32)	195.38 (211.89)	4.95	4.50	4.19
s_y	271.64	271.64	271.64	217.70	217.70	217.70	5.34	5.34	5.34
P.R. Var. Dep. sample	35.96	48.65	49.84	32.57	42.91 (43.07)	49.32 (50.51)	27.01	40.05	45.17
P.R. Var. Ind. sample	32.85	36.18	33.74	27.03	27.53 (18.96)	5.26 (19.45)	14.24	29.13	38.35
Shrinkage (%)	3.11	12.47	16.10	5.54	15.38 (23.11)	44.06 (31.06)	12.77	10.82	6.72

Daily Series Synoptic Weather Maps at sea level and 500 mb. Since these charts are small in their physical size, an estimated rms error in the location of a sea-level anticyclone of 45 n mi appears to be reasonable. This conclusion implies an rms error in displacement of 63.6 n mi.

Finally, there is a question of the accuracy of meteorological analyses both at the surface and at 500 mb. For the last mentioned level, one can perhaps infer from comments by Muench,[4] that a contour analysis rms error of 50 ft is not unreasonable. In addition, linear interpolation between isohypses would introduce an additional error field. Data-analysis errors are extremely important as far as the dependent variables are concerned. For example, since the 24-hr change of central pressure is nearly always a *small* difference between two such error-affected quantities, P' may be seriously affected by "noise." In summary, the dependent variables N', E' and P' are to some extent noise-contaminated.

7. Discussion and summary

Table 6 presents a summary of tests of the various stepwise regression equations generated by the procedures described in earlier sections. For both the south-north (N') and west-east (E') displacements, criterion I, (essentially Miller's criterion) gave by far the smallest shrinkage of the *P.R.* variance. Cutoff criterion II permitted the introduction of 6 variables in each of the prediction equations for N', E' and P', while cutoff III generated prediction equations of even greater length (see Tables 3, 4 and 5). For the prediction of N' and E', increasing the number of predictors (from three) to six or more caused very little increase in *P.R.* variance of the independent sample. On the other hand, the dependent sample tends to be "overfitted," and as a result the shrinkage becomes unrealistically large. In this connection, we quote as a possible explanation a statement due to Panofsky and Brier (1958):

> The equation with the most predictors will not necessarily yield the best fit to the second sample. The reason is that the longest equation may have actually overfitted the first sample and ascribed some of the variation due to small-scale fluctuations to one of the predictors "by accident."

It is of interest to consider meteorologically why Miller's rather stringent cutoff criterion gives prediction equations for N' and E' with the least overfit. Tables 3 and 4a list both the predictors and coefficients involved in these prediction equations, and the predictors are schematically indicated in Fig. 6. If we formulate the system I prediction equations from Tables 3 and 4a, respectively, we may express the results in the form

$$N' = -2.0648(z_{34}-z_{74})+13.8002\bar{T}_{56} -1.3373z_{34}+1225.8887 \quad (6)$$

$$E' = 0.9332(z_{67}-z_{52})+1.3057(z_{65}-z_{52}) +0.7999z_{65}-422.0655 \quad (7)$$

after some rearrangement of terms. Eq (6) and (7) indicate that N' and E' express the commonly accepted synoptic principle of steering of the surface high by the 500-mb contours (see George, 1960, pp. 156–7).

First, consider N' by Eq (6). Normally $z_{34}>z_{74}$ so that the term in parentheses gives a negative or southward displacement. However such a combination of typical values for z_{34} and z_{74} corresponds to a southward-directed geostrophic wind over the sea-level anticyclone. The combined value of the terms in (6) not contained in parentheses will normally be close to zero or negative. It will be helpful to consider the following values to be typical: $z_{34}=800$ ft and $\bar{T}_{56}=-15$C, with the units as indicated in Table 2. The steering principle as expressed by George is further exemplified

[4] Muench, H. S., 1958: Analysis of synoptic charts in the stratosphere, in Contributions to Stratospheric Meteorology. (GRD Research Notes, No. 1), ed. G. Ohring, Air Force Cambridge Research Center.

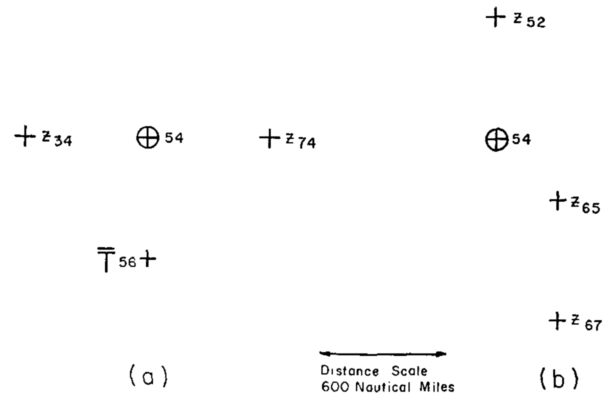

FIG. 6. The set of predictors and grid-point locations relative to a sea-level anticyclone centered at grid-point 54: (a) for N' and (b) for E', respectively, using prediction system I.

by the contribution of $\bar{T}_{56}$: when the surface high is relatively cold (warm), a greater (smaller) southward displacement is likely to be predicted. In this connection, actual observations of N' from the independent sample indicate that when $\bar{T}_{56}$ was much greater than its sample mean, N' was usually numerically small but negative. This is the type of result to be expected from Eq (6) when a warm high ($z_{34} \approx z_{74}$) exists.

Eq (7) for E' indicates a dependence upon steering by a *quasi*-zonal wind. The latter is implied by the positive differences of the form $(z_{67}-z_{52})$ and $(z_{65}-z_{52})$ contained in (7). However the final two terms in (7), which are *not* in parentheses, normally give a positive resultant, even for z_{65} as low as 600 ft. In the case of a *warm anticyclone*, this resultant will be rather large, whereas those in parentheses are small. This indicates that even warm anticyclones are predicted to move sizeable distances to the east, whereas post-mortems of such cases indicated $E' \approx 0$, even though N' tended to be slightly negative. The conclusion to which these results lead is that a *stratification* of the anticyclones into at least "cold" and "warm" types should have been made, with separate regression equations for each type. While Veigas and Ostby did not stratify their cyclones explicitly, they may have accomplished a greater measure of stratification in their study of winter cyclones in the eastern part of the U. S., since this is a preferred region for cyclogenesis. Had the present study been restricted to initially cold anticyclones, greater predictability might have been achieved.

A somewhat surprising result occurred in connection with the prediction equations for P' (Tables 5 and 6). Examination of the results summarized in Table 6 indicates a steady decrease in the *shrinkage* of *P.R.* variance as one proceeds through cutoffs I and II to III, the last of which corresponds to an F-cutoff of four. This meant that the greatest predictability occurred with the largest number of predictors: nine for case III of Table 5. In order to interpret this result, we again examined Sutcliffe's approach (1947, 1950) for a *rationale*. If we use his theory of *intensification*, it is clear that one needs a combination of upper-level and sea-level predictors. Of the nine predictors applicable to P' in system III of Table 5, four apply to upper levels while five are sea-level variables. In addition, five of the nine variables involve 24-hr time-changes either at sea level or at 500 mb, and four are "instantaneous" values from the last available charts prior to the 24-hr forecast. One may speculate therefore that the 24-hr change variables contain the essence of the sea level pressure-change field over the anticyclone, while the other four variables (z_{52}, z_{74}, p_{25}, p_{65}) contain the essence of a steering field for the pressure change field.

Mention must finally be made of our investigations making use of an F-cutout other than zero (see Tables 4b and 6). Actually we made only a very modest effort in this direction, using an F-level change (from zero) of four for cutout, just as we essentially tried similar increments in the F-cutoff. The statistical results obtained for the dependent sample with an F-cutoff of 12.5, indicate that the use of an F-cutout of four causes *no difference* in predictor-selection over that using $F=0$. In fact, when all the predictands were tested using the higher F-cutout, the only one affected was E'. However for E' of Table 6, *both* cutoff and cutout combinations under II and III gave intolerably large shrinkages. Our conclusion as far as this study is concerned is that for prediction of displacements of anticyclones one must use *quite stringent* selection criteria: the most effective selection criteria for anticyclone displacement corresponded to that of Miller (1962). However for the prediction of P', the less stringent criterion F-cutoff$=4.0$ gave the best network of predictor-variables. The best set of prediction equations applied to the sample of 50 cases of independent data gave 24-hr forecasts having a vector rms error of 290 n mi and a central pressure rms error of 4.2 mb. These results are reasonably compatible with the accuracy of those of Veigas and Ostby, despite the fact that anticyclones are susceptible to greater errors, both in location and in central pressure, than cyclones.

In connection with the selection methods employed in this paper, improvement may be possible if there should exist an objective technique for determining an optimum relationship between F-cutoff and F-cutout. There seems to be very little conclusive statistical evidence concerning such an optimum relationship at the present time.

Acknowledgments. The authors wish to thank Professor W. D. Duthie of the U. S. Naval Postgraduate School for helpful and stimulating discussions during the preparation of this paper. Also our appreciation is due to the Computing Center of the U. S. Naval Postgraduate School for their preparation of the

BIMD 09 program and for their operation of the computer.

REFERENCES

Anderson, T. W., 1960: *An introduction to multivariate statistical analysis.* New York, London, John Wiley and Sons, 374 pp.

Efroymson, M. A., 1960: Multiple regression analysis, in *Mathematical methods for digital computers.* New York, John Wiley and Sons, pp. 191–203.

George, J. J., 1960: *Weather forecasting for aeronautics.* New York and London, Academic Press, 673 pp.

Haltiner, G. J., and F. L. Martin, 1957: *Dynamical and physical meteorology.* New York, Toronto, London, McGraw-Hill Book Company, Inc., 470 pp.

Miller, R. G., 1962: Statistical prediction by discriminant analysis. *Meteor. Monogr.*, 4, No. 25, Boston, Mass., Amer. Meteor. Soc., 53 pp.

Panofsky, H. A., and G. W. Brier, 1958: *Some applications of statistics to meteorology.* Pennsylvania State Univ., University Park, Pa., 224 pp.

Sutcliffe, R. C., 1947: A contribution to the problem of development. *Quart. J. R. meteor. Soc.*, **73**, 370–383.

——, and A. G. Forsdyke, 1950: The theory and use of upper-air thickness patterns in forecasting. *Quart. J. R. meteor. Soc.*, **76**, 189–217.

Veigas, K. W., and F. P. Ostby, 1963: Application of a moving-coordinate prediction model to East Coast cyclones. *J. appl. Meteor.*, **2**, 24–38.